

# Memorandum

**TO:** Commissioners  
**FROM:** Paul J. Hofer, Senior Research Associate  
**DATE:** June 1, 1998  
**REGARDING:** Update on Disparity Research

Several interesting developments have occurred in the past year and a half in the field of federal sentencing disparity research. Staff have made considerable progress on our in-house research; two papers were completed, reviewed by our expert Advisory Panel, and are ready for publication. In addition, several papers by outside researchers have been published, presented at professional conferences, or distributed to the press. The time is right to update the Commission on what we know about sentencing disparity under the guidelines.

The recent research can be divided into two topics: 1) disparity created by differences among judges, and 2) disparity related to the characteristics of the offender, such as gender or race. We will discuss developments in each of these areas in turn.

The Commission's work on disparity has been undertaken with the assistance of an outside panel of experts, the Disparity Research Advisory Panel, chaired by Professor John Hagan of the University of Toronto. Professor Hagan and other panel members have reviewed research proposals for all major disparity-related projects, including those of Paul Hofer, Kevin Blackwell, and Phyllis Newton. In addition, the panel as a whole reviewed a draft report of Hofer and Blackwell's inter-judge disparity study. Professor Hagan also reviewed the final version of that report and the methodological paper on discrimination research discussed below. Professor Hagan's most recent comments on the Commission's in-house work are attached.

## Research on Inter-judge Sentencing Disparity

***In-house studies.*** In November 1996 Kevin Blackwell and I briefed Commissioners on our study that analyzed the effect of the guidelines on inter-judge disparity. This study was subsequently reviewed by our Advisory Panel, which recommended several additional analyses. We implemented as many of these as were technically feasible and sent the revised paper to Professor Hagan for review. He now recommends submission of the paper to the JOURNAL OF CRIMINAL LAW AND CRIMINOLOGY, which he serves as Professional Criminology Editor.

The principle findings of this study can be summarized briefly:

- Considering the nation as a whole and the overall criminal docket, the guidelines have reduced sentencing disparity due to differences among judges. The amount of variation due to judge assignment dropped by almost half between 1984-85 and 1994-95. (What this decrease means in actual months in prison is discussed in the next section.) This extends the finding from the Commission's Four-Year Evaluation, which found a

reduction in the dispersion of sentences for several groups of matched cases.

- The success of the guidelines at reducing disparity is uneven across different regions and different types of offenses. Most notably, there is evidence that regional differences in the sentencing of drug cases may have *increased* under the guidelines. This is partly due to the greater length of drug sentences, which makes variations among regions more pronounced. Further research is needed to identify the sources of this disparity. Variations in how different districts use departures are one possible source;<sup>1</sup> differences in plea bargaining practices are another.

The methodology used for this study was the “natural experiment” created by the random assignment of cases to judges, which is used in most districts. Social scientists consider randomized research designs the most powerful evaluation tools available and considerable statistical sophistication is required. We believe this paper will help establish the Commission as a center for top-flight sentencing research.

**External research.** The elegance of the natural experiment methodology has attracted outside researchers; two papers appeared within the last year evaluating the guidelines using this approach. The first was published by Professor A. Payne in the *INTERNATIONAL REVIEW OF LAW AND ECONOMICS*.<sup>2</sup> Like us, Payne found that the guidelines reduced inter-judge disparity in some, but not all, districts and for some, but not all, types of cases. Our paper now discusses Payne’s study and several methodological limitations upon which we were able to improve. For example, our study used more recent data and included a greater number of districts.

In December, 1997, James Anderson, Jeffrey Kling, and Professor Kate Stith of Yale Law School presented a paper using the natural experiment approach at a conference sponsored by the *JOURNAL OF LAW AND ECONOMICS*. Mr. Kling, a statistician with the Massachusetts Institute of Technology, developed a powerful statistical model that measured changes in the amount of inter-judge disparity from 1983 through 1992. Similar to our study, these researchers’ findings — though preliminary and not ready for citation — suggest that by 1992 the guidelines cut the amount of this type of disparity by three-quarters. Using the “Gini coefficient” as a measure of differences among judges, they translated this effect into absolute months of imprisonment. From 1986-87 to 1991-92, the amount of variation in sentence length due to the judge to which a case was assigned fell from an average of 8.5 months to an average of 3.1 months.

---

<sup>1</sup> For a recently-appearing comparison of two contiguous districts that adapted to the guidelines differently, which could lead to the type of regional differences we found, see Lisa M. Farabee, “Disparate Departures Under the Federal Sentencing Guidelines: A Tale of Two Districts,” 30 *CONN. L. REV.* 569 (1998).

<sup>2</sup> A. Abigail Payne, “Does Inter-Judge Disparity Really Matter? An Analysis of the Effects of Sentencing Reforms in Three Federal District Courts,” 17 *INT’L REV. OF L & ECON.* 337 (1997).

Together, these studies provide powerful evidence — the most powerful evidence likely to be developed — that the guidelines have successfully achieved one of their primary goals: reducing unwarranted disparity. These studies also support a long line of sentencing research concluding that the greatest cause of unwarranted disparity is not discrimination against particular groups, but philosophical differences among judges.

## **Research on Disparity Due to Characteristics of the Offender**

**External research.** The most popular topic of disparity research, however, continues to be racial and gender disparity. Differences in average sentences for Blacks, Whites, men, and women can easily be demonstrated. But researchers are not always careful to define how much, if any, of these differences are unwarranted and how much, if any, may reflect discrimination against a particular group. Much of the work in this area is also plagued by methodological problems that the researchers have not fully appreciated.

Two studies appeared in the past year that purport to demonstrate the influence of bias or racial and gender stereotypes. Albonetti, writing in the *LAW & SOCIETY REVIEW*, reported “nontrivial” effects for both gender and race.<sup>3</sup> For example, after controlling for offense level, criminal history score, and several other factors, women were six percent less likely than men to be imprisoned and Blacks were three percent more likely than Whites.

In late April, the Commission received calls from reporters about an unpublished paper by Professor David Mustard of the University of Georgia.<sup>4</sup> It concluded that race, ethnic, and gender-based disparities are “prevalent” in federal sentencing and that legally-relevant factors do not appear to account for all the differences observed. For example, among non-departure cases, the average sentence for Blacks was found to be about two months longer at the mean than the average sentence for Whites.

Kevin and I reviewed a draft of this paper and wrote a response, which we sent to Professor Mustard and would happily share with anyone else who is interested. Based on our suggestions, Professor Mustard decided to redo his analyses for publication at a later date. This study had one of the same problems as the study published in the *Nashville Tennessean* several years ago: it failed to control for the effects of mandatory minimum statutes that truncate or trump the guideline range. In a re-analysis of the data used in the *Tennessean* study, Kevin found that the mandatory minimums completely accounted for the race effect.

---

<sup>3</sup> Celeste A. Albonetti, “Sentencing under the Federal Sentencing Guidelines: Effects of Defendant Characteristics, Guilty Pleas, and Departures on Sentence Outcomes for Drug Offenses, 1991-1992,” 31 *LAW & SOC REV* 789 (1997).

<sup>4</sup> David B. Mustard, “Racial, Ethnic and Gender Disparities in Sentencing: Evidence from the US Courts,” Revision February 1998. (Unpublished paper on file with the author.)

***Internal research.*** Examining the dozen studies of racial and gender disparity under the federal guidelines (as well as scores of studies on sentencing in other jurisdictions) we concluded that many of the same problems are common to most of them. This prompted us to write a methodological review and critique of research on discrimination, which we presented at the annual meeting of the American Criminological Association.<sup>5</sup> The paper received a favorable review from Dr. William Rhodes of Abt Associates, the discussant for the panel at the meeting. It was subsequently reviewed by Professor Hagan, who recommends minor revisions and submission for publication in *LAW & SOCIETY REVIEW* or *JUSTICE QUARTERLY*. (See attachment for Professor Hagan's recommendations.)

Among the points raised in the paper are:

- Many of the findings in current studies of sentencing disparity reflect methodological problems and do not demonstrate discrimination on the part of judges.
- Researchers should not infer discrimination unless they have controlled for the most important legally-relevant factors that may account for differences among groups. These factors are specific to the particular decision being made.
- The most significant reason for differences among average sentences for Blacks and Whites is not discrimination. It is the effects of factors that are legally relevant, but that have a disproportionate impact on Blacks, such as the harsher treatment of crack cocaine.

We believe publication of this paper in a widely-circulated professional journal will prevent future researchers from making the same mistakes. It may also help redirect resources toward more fruitful lines of inquiry.

---

<sup>5</sup> Paul J. Hofer & Kevin R. Blackwell, "Identifying Sources of Unfairness in Federal Sentencing," presented November 21, 1997, at the convention of the American Criminological Association. (On file with the author.)